شبکه GAN چیست؟ معرفی شبکه مولد تخاصمی به زبان ساده

شبکه مولد تخاصمی یا GAN چیست؟ مولد یا generator در GAN چه کاری انجام می‌دهد؟ متمایزگر یا discriminator چیست؟ GANها چگونه کار می‌کنند؟ مراحل آموزش GAN‌ها چگونه است؟ GANها چه کاربردهایی دارند؟ انواع مختلف شبکه های GAN  کدامند؟

[یادگیری عمیق](https://hamruyesh.com/deep-learning-neural-net-history-guide-part-4/) (Deep Learning) کاربردهای بی شماری در صنایع پیدا کرده است. [شبکه‌های عصبی](https://hamruyesh.com/product/artificial-neural-networks-tutorial-from-scratch/) تقریباً در همه بخش‌های تجاری مانند مراقبت‌های بهداشتی، امور مالی و تجارت الکترونیک به طور گسترده مورد استفاده قرار می‌گیرند. این [شبکه‌های عصبی](https://hamruyesh.com/deep-learning-neural-net-history-guide-part-3/) (neural networks) به حل مشکلات کسب و کار کمک می‌کنند.

در این مقاله از مجله [هم‌رویش](https://hamruyesh.com/)، با شبکه‌های مولد تخاصمی یا GAN‌ها آشنا می‌شویم و به سوالات مطرح شده پاسخ می‌دهیم.

**Gan چیست ؟ (شبکه مولد تخاصمی)**

شبکه مولد تخاصمی (Generative Adversarial Network) یا به اختصار GAN، در سال 2014 توسط یان گودفلو (Ian J. Goodfellow) و نویسندگان همکار او معرفی شد. GAN‌ها وظایف یادگیری بدون نظارت را در [یادگیری ماشین](https://hamruyesh.com/deep-learning-vs-machine-learning-a-simple-explanation/) انجام می‌دهند. این شبکه‌ها شامل دو مدل هستند که به طور خودکار، الگوهای موجود در داده‌های ورودی را کشف کرده و یاد می‌گیرند. این دو مدل با نام‌های مولد (Generator) و متمایزگر (Discriminator) شناخته می‌شوند.

 مولد و متمایزگر، برای بررسی، ضبط و تکرار تغییرات درون مجموعه داده با یکدیگر رقابت می‌کنند. می‌توان از GAN‌ها برای تولید نمونه‌های جدیدی که به طرز قابل قبولی از مجموعه داده اصلی قابل تهیه هستند، استفاده کرد.

در شکل زیر، نمونه‌ای از GAN نشان داده شده است. یک پایگاه داده حاوی اسکناس‌های واقعی 100 روپیه‌ای وجود دارد. [شبکه عصبی](https://hamruyesh.com/product/artificial-neural-networks-tutorial-from-scratch/) مولد، اسکناس‌های جعلی 100 روپیه‌ای را تولید می‌کند. شبکه متمایزگر، به شناسایی اسکناس‌های واقعی و جعلی کمک می‌کند.



**مولد چیست؟**

یک [شبکه عصبی](https://hamruyesh.com/product/artificial-neural-networks-tutorial-from-scratch/) است که داده‌های جعلی تولید می‌کند تا متمایزگر توسط آن‌ها آموزش ببیند. مولد یاد می‌گیرد که داده‌های قابل قبول تولید کند. مثال ها/نمونه‌های تولید شده، برای متمایزگر، نمونه‌های منفی آموزشی به حساب می‌آیند. مولد، یک بردار نویز تصادفی با طول ثابت را به عنوان ورودی می‌گیرد و یک نمونه تولید می‌کند.

generator چیست؟

هدف اصلی مولد این است که متمایزگر را طوری فریب دهد که خروجی خود را با عنوان “واقعی” دسته بندی کند. قسمتی از GAN که مولد را آموزش می‌دهد شامل موارد زیر است:

* بردار ورودی نویزی
* شبکه مولد، که ورودی تصادفی را به یک نمونه داده تبدیل می‌کند
* شبکه متمایزگر، که داده‌های تولید شده را دسته بندی می‌کند
* تلفات مولد، که مولد را به دلیل احمق پنداشتن متمایزگر، مجازات می‌کند!

از روش پس انتشار (backpropagation) برای تنظیم هر وزن در جهت مناسب با محاسبه تاثیر وزن بر خروجی استفاده می‌شود. همچنین از این روش برای به دست آوردن گرادیان استفاده می‌شود و این گرادیان‌ها می‌توانند به تغییر وزن‌های مولد کمک کنند.



بیایید به سراغ موضوع بعدی این مقاله برویم، یعنی متمایزگر.

**متمایزگر چیست؟**

متمایزگر (Discriminator) یک [شبکه عصبی](https://hamruyesh.com/product/artificial-neural-networks-tutorial-from-scratch/) است که داده‌های واقعی را از داده‌های جعلی ایجاد شده توسط مولد تشخیص می‌دهد. داده‌های آموزشی متمایزگر از دو منبع مختلف به دست می‌آیند:

* نمونه‌های واقعی داده‌ها، مانند تصاویر واقعی پرندگان، انسان‌ها، اسکناس‌های ارزی و غیره، توسط متمایزگر به عنوان نمونه‌های مثبت در طول آموزش استفاده می‌شوند.
* نمونه‌های جعلی ایجاد شده توسط مولد، به عنوان نمونه‌های منفی در طول فرایند آموزش استفاده می‌شوند.

Discriminator چیست؟

در حین آموزش، متمایزگر به دو تابع تلفات (loss function) متصل است. در طول آموزش شبکه متمایزگر، از تلفات مولد چشم پوشی شده و فقط از تلفات متمایزگر استفاده می‌شود.

متمایزگر، در حین فرآیند آموزش، داده‌های واقعی و داده‌های جعلی دریافتی از مولد را دسته بندی می‌کند. در صورت دسته بندی نادرست یک نمونه داده واقعی به عنوان نمونه جعلی یا برعکس، توسط تلفات متمایزگر مجازات می‌شود.

متمایزگر، وزن‌های خود را با پس انتشار از تلفات متمایزگر در طول شبکه خود، به روز می‌کند.



اکنون با نحوه عملکرد GAN‌ها در این مقاله آشنا می‌شویم.

**GAN‌ها چگونه کار می‌کنند؟**

GAN از دو [شبکه عصبی](https://hamruyesh.com/deep-learning-neural-net-history-guide-part-3/) تشکیل شده است. یک مولد G(x) و یک متمایزگر D(x). هر دوی آن‌ها یک بازی خصمانه (adversarial) انجام می‌دهند. هدف مولد این است که با تولید داده‌هایی مشابه داده‌های موجود در مجموعه آموزشی، متمایزگر را فریب دهد. متمایزگر سعی خواهد کرد با شناسایی داده‌های جعلی از داده‌های واقعی فریب نخورد. هر دو به طور همزمان برای یادگیری و آموزش داده‌های پیچیده مانند فایل‌های صوتی، ویدئویی یا تصویری کار می‌کنند.

شبکه مولد یک نمونه را می‌گیرد و یک نمونه جعلی از داده‌ها را تولید می‌کند. مولد آموزش دیده است تا احتمال بروز اشتباه توسط شبکه متمایزگر را افزایش دهد.



در شکل زیر، نمونه‌ای از GAN نشان داده شده است که سعی می‌کند واقعی یا جعلی بودن اسکناس‌های 100 روپیه‌ای را تشخیص دهد. بنابراین، ابتدا یک بردار نویز یا بردار ورودی به شبکه مولد تغذیه می‌شود. مولد، اسکناس‌های جعلی 100 روپیه‌ای تولید می‌کند. تصاویر واقعی اسکناس‌های 100 روپیه‌ای که در یک پایگاه داده ذخیره شده، به همراه اسکناس‌های جعلی، به متمایزگر منتقل می‌شوند. در ادامه، متمایزگر، اسکناس‌های واقعی را از اسکناس‌های جعلی متمایز و دسته بندی می‌کند.

ما مدل را آموزش می‌دهیم، تابع تلفات را در انتهای شبکه متمایزگر محاسبه می‌کنیم و این تلفات را از طریق پس انتشار به دو مدل متمایزگر و مولد ارائه می‌دهیم.



**مراحل آموزش GAN**

* تعریف مسئله
* انتخاب معماری GAN
* آموزش متمایزگر با داده‌های واقعی
* تولید ورودی‌های جعلی توسط مولد
* آموزش متمایزگر با داده‌های جعلی
* آموزش مولد با خروجی متمایزگر

اکنون اجازه دهید تا با انواع شبکه های gan آشنا شویم.

**انواع شبکه های gan**

**GANهای وانیلی:**

Vanilla GANs دارای یک فرمول بهینه سازی حداقل حداکثر (min-max) هستند که در آن، متمایزگر یک دسته بندی باینری انجام می‌دهد و از تلفات آنتروپی متقابل سیگموئید (sigmoid cross-entropy) هنگام بهینه سازی استفاده می‌کند. مولد و متمایزگر در GANهای وانیلی، پرسپترون‌های چند لایه هستند. این الگوریتم سعی می‌کند معادله ریاضی حاکم بر GANها را با استفاده از گرادیان نزولی تصادفی بهینه کند.

**GANهای کانولوشنی عمیق:**

Deep Convolutional GANs یا به اختصار DCGANs، از [شبکه‌های عصبی کانولوشنی](https://hamruyesh.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/) به جای [شبکه‌های عصبی](https://hamruyesh.com/product/artificial-neural-networks-tutorial-from-scratch/) وانیلی در مولد و متمایزگر پشتیبانی می‌کنند. این شبکه‌ها از پایداری بیشتری برخوردار هستند و تصاویری با کیفیت بهتر تولید می‌کنند.

مولد، مجموعه‌ای از لایه‌های کانولوشنی با کانولوشن‌های دارای گام کسری (fractional-strided) یا کانولوشن¬های ترانهاده‌ای (transpose) است، بنابراین تصویر ورودی را در هر لایه کانولوشنی، رو به بالا نمونه برداری (up-sample) می‌کند.

متمایزگر، مجموعه‌ای از لایه‌های کانولوشنی با کانولوشن‌های دارای گام (strided) است، بنابراین تصویر ورودی را در هر لایه کانولوشنی، رو به پایین نمونه برداری (down-sample) می‌کند.

 **GANهای شرطی:**

GAN‌های وانیلی را می‌توان با استفاده از اطلاعات اضافی و برای تولید نتایج بهتر، به مدل‌های شرطی بسط داد. در Conditional GAN یا به اختصار CGAN، یک پارامتر اضافی “y” برای تولید داده‌های مربوطه به مولد اضافه می‌شود. برچسب‌ها به عنوان ورودی به متمایزگر تغذیه می‌شوند تا به تشخیص داده‌های واقعی از داده‌های جعلی تولید شده کمک کنند.

**GANهای با وضوح عالی:**

Super Resolution GANs یا به اختصار SRGANs، از [شبکه‌های عصبی عمیق](https://hamruyesh.com/deep-learning-neural-net-history-guide-part-2/) به همراه یک شبکه تخاصمی برای تولید تصاویر با وضوح بالاتر استفاده می‌کنند. SRGAN‌ها هنگامی که تصویری با وضوح پایین به آن‌ها داده می‌شود، تصویری واقع گرایانه با رزولوشن بالا تولید می‌کنند.

حالا که با GAN‌ها آشنا شدیم، بیایید نگاهی به برخی کاربردهای مهم GAN‌ها بیندازیم.

**کاربردهای GAN**

* با کمک DCGAN‌ها می‌توانید تصاویر شخصیت‌های کارتونی را به منظور تولید چهره شخصیت‌های انیمه‌ای و همچنین شخصیت‌های پوکمون (Pokemon) آموزش دهید.

کاربرد 1

* GAN‌ها را می‌توان بر روی تصاویر انسان‌ها آموزش داد تا چهره‌هایی واقع گرایانه تولید کنند. چهره‌هایی که در تصویر زیر مشاهده می‌کنید با استفاده از GANها ایجاد شده‌اند و در واقعیت وجود ندارند.

کاربرد 2

* GAN‌ها می‌توانند تصاویری واقع گرایانه از توصیف متنی اشیائی نظیر انسان‌ها، پرندگان و دیگر حیوانات بسازند. ما یک جمله به عنوان ورودی به شبکه می‌دهیم و چندین تصویر متناسب با آن توصیف تولید می‌کنیم.

در شکل زیر، نمونه‌ای از ترجمه متن به تصویر با استفاده از GANها برای پرنده‌ای با سر سیاه، بدن زرد و بادخور کوتاه آورده شده است.

کاربرد 3

**نتیجه گیری**

پس از خواندن این مقاله، شما با اصول GAN و نحوه کار آن آشنا شده‌اید. در نهایت مفهوم مولد و متمایزگر را درک کردید. شما دیدید که چگونه GAN‌ها می‌توانند به شناسایی تصاویر واقعی و جعلی کمک کنند. همچنین فرآیند آموزش شبکه مولد تخاصمی و کاربردهای GAN‌ها را بررسی کردید.